สหัส พรหมสิทธิ์

ชุบชีวิตด้วยคอมพิวเตอร์ไฮเทค

การผสมผสานความรู้ด้าน computer science เข้ากับวิทยาศาสตร์ชีวภาพ มาไขความเร้นลับในโลกของสิ่งมีชีวิต

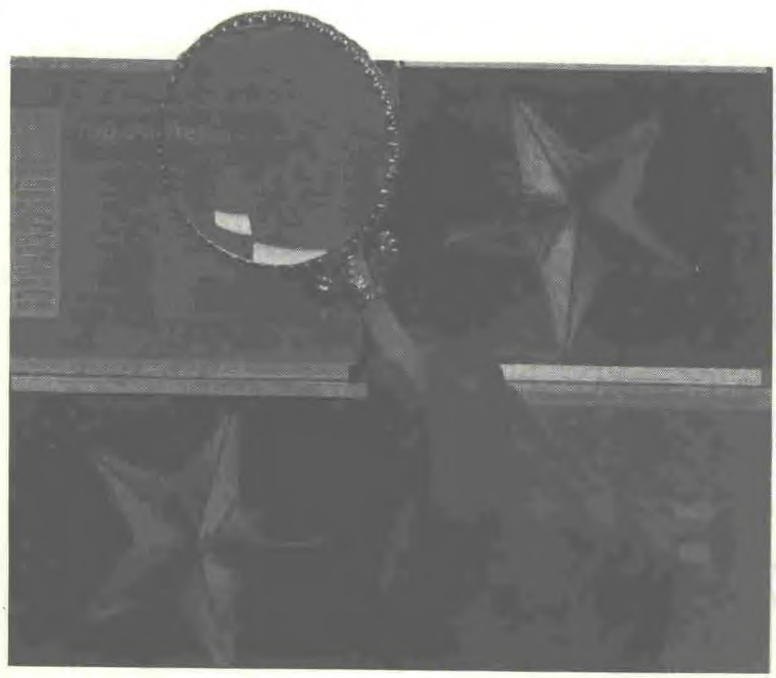
ไม่น่าจะมีความยิ่งใหญ่อันใดในโลกของคอมพิวเตอร์เท่ากับการที่มีผู้พยายามผสมผสานความรู้ในด้าน computer science ให้เข้ากับวิทยาศาสตร์ชีวภาพจนกลายมาเป็นวิชาแขนงใหม่ที่เรียกว่า computational biology อันเป็นการนำคอมพิวเตอร์ศาสตร์มาไขความเร้นลับในโลกของสิ่งที่มีชีวิต (เช่นในตัวมนุษย์) โดยมองจากทัศนะของ molecular biology

การที่จะเข้าถึงเรื่องนี้ให้ได้อย่างกระจ่างชัดคงหนีไม่พ้นการตามไปดูงานค้นคว้าวิจัยในด้านชีววิทยาภาคคำนวณ (computational biology) ที่กำลังดำเนินอยู่ในมหาวิทยาลัยต่าง ๆ ทั้งในสหรัฐฯ และประเทศอื่น ว่าได้มีการทุ่มเทค้นคว้ากันอย่างไรบ้าง และคอมพิวเตอร์ได้เข้ามามีบทบาทแค่ไหนในทางนี้ รวมทั้งมีอะไรบ้างที่เป็นหัวใจสำคัญที่นักวิทยาศาสตร์และนักคอมพิวเตอร์ศาสตร์ต้องขบให้แตกเพื่อก้าวไปสู่ยุคของ**การปฏิรูปวิวัฒนาการ** (euolution revolution) อันเป็นยุคที่มนุษย์สามารถกำชะตาชีวิตของตนเองได้โดยไม่ต้องตกอยู่ภายใต้ทฤษฎีวิวัฒนาการที่สืบทอดไปในมิติของเวลาที่นานนับล้านปี อันหมายถึงว่าในต่อ ๆ ไปนั้นมนุษย์สามารถอาศัยวิทยาศาสตร์และเทคโนโลยีด้านคอมพิวเตอร์เป็นเครื่องมือในการสร้างมนุษย์ได้ ไม่ว่าจะเป็นมนุษย์แท้หรือมนุษย์เทียม พร้อมทั้งสามารถกำหนดในขั้นรายละเอียดว่าต้องการมนุษย์ชนิดไหน ความรู้ในด้านชีววิทยาในระดับโมเลกุล (molecular biology) จึงเป็นพื้นฐานสำคัญของการก้าวไปสู่ **"การปฏิรูปวิวัฒนาการ"** ซึ่งคนกลุ่มหนึ่งเคยใฝ่ฝันมานานแล้ว และอาจจะเป็นไปได้จริงในอนาคตอันใกล้นี้ เนื่องจากการมีคอมพิวเตอร์ที่มีอานุภาพมาช่วยในการค้นคว้าเพื่อให้ความฝันเป็นความจริง

Human Genome Project (1990–2005)

ในแง่ของความซับซ้อนที่สุดในโลกคงไม่มีอะไรเกินสิ่งมีชีวิตและจักรกลอันสำคัญที่กุมความลับของสิ่งมีชีวิตเช่นมนุษย์ ซึ่งคงหนีไม่พ้นเรื่องของ DNA เมื่อเป็นเช่นนี้สิ่งแรกที่ต้องทำความเข้าใจก็คือเรื่องของหน่วยของชีวิตที่มีขนาดเล็กมาก ๆ ที่ถือว่าเป็นเบสิกของเรื่องนี้ก็คือยิ่งรู้ได้ลึกซึ้งมากเท่าไหร่ก็ยิ่งจะเป็นกำไรในการค้นคว้ามากขึ้นเท่านั้น ดังจะเห็นจากผู้

ที่เป็นตัวเอกในการค้นคว้าทางด้านนี้ที่มีทั้งนายแพทย์และนักวิทยาศาสตร์ในระดับปรมาจารย์ ซึ่งยังอยู่ในวัยหนุ่มใหญ่ที่มีอายุอานามราว 40-50 ปี พร้อมด้วยคุณวุฒิอันยาวเหยียด ดังจะพาท่านผู้อ่านไปทำความรู้จักในภายหลัง สิ่งที่อยากจะพูดในขั้นแรกนี้จะเป็นเรื่องของอนุภาคของชีวิต เช่น โมเลกุลที่เป็น DNA ว่ามีความสำคัญอย่างไรบ้างกับชีวิต การรวมตัวระหว่าง computer science กับ human biology ได้ทำให้เกิดเป็นวิชา computational biology นักค้นคว้าที่ทำงานในด้านนี้ จึงได้ชื่อว่าเป็น computational biologists ที่อาศัยเทคโนโลยีด้านคอมพิวเตอร์เข้าช่วย โดยมีสิ่งสนับสนุนขั้นพื้นฐาน เช่น 3-D modeling software, fuzzy-logic search algorithms ซึ่งสามารถไขความเร้นลับว่ายีนส์ (genes) ต่าง ๆ ควบคุมการติดโรคของมนุษย์ได้อย่างไร รวมทั้งการที่ร่างกายมนุษย์ไม่สามารถรักษาโรคให้ตัวเองได้ แต่ถ่ายทอดสิ่งที่เป็นกรรมพันธุ์จากชั่วคนหนึ่งไปยังอีกชั่วคนหนึ่งได้ อันเป็นการถ่ายทอดทั้งสิ่งที่ดีและไม่ดี



พันธุกรรมศาสตร์ (Genetics) เคยเป็นวิชาที่กฎและหลักเกณฑ์หลาย ๆ อย่างยังแทรกเข้าไปไม่ถึง (vague science) จึงน่าจะเป็นศาสตร์ที่ต้องอาศัยความจัดเจน (empirical) เป็นโครงสร้าง อย่างเช่น ทุกคนรู้ว่าอุปนิสัย (traits) หรือคุณสมบัติของคนจะมีการสืบทอดจากชั่วอายุคนหนึ่งไปยังอีกชั่วคนโดยไม่มีใครรู้ว่าสืบทอดอย่างไร นักวิทยาศาสตร์จึงได้แต่เพียงสันนิฐานว่า อุปนิสัยแต่ละอย่างคงจะมีการสืบทอดในรูปของการสื่อสารด้วยตัวแทนที่เรียกว่ายีนส์ (genes) ซึ่งเมื่อรวมกันเข้าแล้วก็จะกลายเป็นสิ่งที่เรียกว่า เจนโนมของมนุษย์ (genome) ที่หมายถึง เจเนติกโค้ด (genetic code) ทั้งหมดที่สำหรับใช้สร้างมนุษย์ขึ้นมา แต่ก็ยังไม่มีใครรู้ว่า เจนโนมทำงานอย่างไรและมีหน้าตาเป็นเช่นใด

ในปี 1953 ได้มีการพบว่าในแต่ละเซลล์ของสิ่งมีชีวิตเช่นมนุษย์จะมีสารที่เรียกว่า deoxyribonncleic acid อยู่ภายในนั้น คือนั่งอยู่ในทุกเซลล์ โดยทำหน้าที่เป็นจักรกลในการควบคุม (control mechanism) เจนโนมของมนุษย์มียีนอยู่ราว 100,000 ตัวที่มีลักษณะเป็น DNA ที่มีความยาวไม่มาก โดยทำหน้าที่บอกวิธีสร้างโปรตีนให้แก่เซลล์ โปรตีนนี้ถือว่าเป็นโครงสร้างขั้นพื้นฐานของชีวิต DNA ชนิดต่าง ๆ จะมีอยู่ในสิ่งมีชีวิตอันนับตั้งแต่แบคทีเรียจนถึงสัตว์น้ำที่มีขนาดอันมหึมาเช่นปลาวาฬ สารเคมีที่เป็นองค์ประกอบของ DNA มีอยู่ 4 อย่างคือ adenine (A), thymine (T), cytosine (C) และ guanine(G) โดยจับกันเป็นคู่ต่าง ๆ ในรูปของ double helix อันดับของการจัดเรียง (sequence) ของเบสแปร์ต่าง ๆ (base pairs) ไม่เพียงแต่จะแปรไปจากสกุล (species) หนึ่งไปยังสกุลถัดไปเท่านั้น แต่ยังแปรตามตัวบุคคลหรือสิ่งมีชีวิตแต่ละตัว อันจะมีผลตั้งแต่อาทิเช่น สีของผม จนถึงความยืนยาวของอายุของมนุษย์หรือของสัตว์ นอกจากนี้ก็มีเรื่องของการผ่าเหล่า (mutations) ที่มีทั้งดีและไม่ดี คือในบางครั้งก็เป็นไปในทางการปรับปรุงให้ดีขึ้นอันเป็นวิวัฒนาการเชิงสร้างสรรค์ แต่ในหลาย ๆ ครั้งก็เป็นไปในทางไม่ดี เช่น ไปแทรกแซงการทำงานที่เป็นปกติดีอยู่แล้วให้เกิดการผิดปกติ ทำให้เป็นที่มาของโรคภัยไข้เจ็บต่าง ๆ ที่มีสมุหฐานมาจากความบกพร่องของ DNA

ในขั้นแรกของการค้นคว้าเพื่อหาทางแก้ไขข้อบกพร่องของ DNA ดูเหมือนว่าน่าจะเป็นของยากจนเป็นไปไม่ได้ เพราะ DNA ที่มีขนาดเท่าโมเลกุลนี้จะเล็กจนยากแก่การใช้อุปกรณ์เครื่องมือธรรมดามาทำการตรวจสอบและแก้ไข อย่างไรก็ตาม ก็ยังมีทางออกสำหรับเรื่องนี้ในการค้นคว้าทดลอง ด้วยการใช้**เอ็นไซม์** (enzymes) บางชนิด ที่สามารถตัด DNA ออกเป็นชิ้นเล็ก ๆ ตรงจุดที่กำหนดไว้ ทำให้สามารถนำส่วนที่ตัดออกมาเป็นเซกชัน (sections) ไปทำการซ่อมแซมหรือปรับปรุงเสร็จแล้วก็มีการเชื่อมกลับเข้าที่ดังเดิมอันเป็นที่มาของคำว่า **การเชื่อมต่อยีน** (gene splieing) ที่เป็นไปได้ในทางปฏิบัติในเมื่อไม่กี่ปีมานี้

จากนี้ก็จะขอย้อนกลับไปในเรื่องของบทบาทของคอมพิวเตอร์ที่เข้ามาสัมพันธ์กับการฝืนวิวัฒนาการ หรือการเอาชนะธรรมชาตินับตั้งแต่เรื่องของการเจ็บไข้ได้ป่วย จนถึงแนวการค้นคว้าอื่น ๆ เพื่อปรับปรุงชีวิตให้ดียิ่งขึ้นในรูปของคอมพิวเตอร์ศาสตร์ที่ผสมผสานกับชีววิทยาด้านมนุษย์ โดยจะเริ่มที่**โครงการเจนโนมเพื่อมนุษยชาติ** (Human Genome Project) อันเป็นโครงการการค้นคว้าระดับโลกซึ่งได้มีการก่อตั้งมาหลายปีแล้ว และคาดว่าในราวปี ค.ศ. 2001 ถึง 2005 โครงการนี้จะก้าวไปสู่ความสำเร็จครั้งยิ่งใหญ่ในการสร้างโค้ดที่สมบูรณ์สำหรับควบคุมการสร้างชีวิตมนุษย์ (to transcribe the code that controls the

creation of human life) ด้วยการใช้เงินถึง **3,000 ล้านเหรียญสหรัฐฯ** โดยมี National Institute of Health (NIH) ซึ่งตั้งอยู่ในรัฐแมรีแลนด์เป็นผู้ประสานงานกับศูนย์วิจัยที่อยู่ในประเทศอื่น (เช่น ญี่ปุ่น) คือเมื่อถึงเวลานั้นอันดับแรกของโมเลกุลของ DNA ของมนุษย์ก็จะเข้าไปอยู่ในรูปของแถวอันยาวเหยียดของบิต (bits) ซึ่งจะกินเนื้อที่ราว **750 เมกะไบต์** ในฮาร์ดไดรฟ์ สำหรับในกรณีที่ใช้ data compression เข้าช่วย ความยาวดังกล่าวก็จะสั้นกว่า 750 เมกฯ โดยลดลงไปได้มาก ซึ่งในขณะนี้ก็ได้มีโค้ดเช่นที่ว่านี้แจกจ่ายให้ฟรีแล้ว สำหรับผู้ที่มีโมเด็มและ browser ซอฟต์แวร์ ซึ่งจะสามารถดาวน์โหลด (down load) โปรแกรมนี้จาก GenBank Web site ของโครงการนี้ (Human Genome Project) แต่ก็ต้องเข้าใจว่าโค้ดเช่นที่ว่านี้เป็นเพียงงานขั้นต้นที่ยังขาดความสมบูรณ์ที่มีไว้ให้ผู้สนใจนำไปศึกษาหาความรู้ด้วยตนเอง ซึ่งจะต้องรออีกสักพักกว่าจะก้าวไปถึงจุดนั้น คือเมื่อโลกได้ก้าวไปถึงจุดที่สามารถถอดรหัสของ DNA ของมนุษย์ได้หมด พร้อมทั้งได้มีการทำเครื่องหมายกำกับไว้ในจุดต่าง ๆ ตลอดความยาวของ helix ของ DNA แล้วแปลงข้อมูลที่แฝงอยู่ในสาย DNA ให้เป็นโค้ด (code) ทางคอมพิวเตอร์ เมื่อนั้นมนุษย์ก็จะสามารถนำโค้ดที่ว่านี้ไป recompile เพื่อสนองประโยชน์หรือสนองความสนใจเป็นการส่วนตัวได้ในประการต่าง ๆ อันไม่ว่าจะเป็นการสร้างคนขึ้นมาให้เป็นไปตามความต้องการหรือนำโค้ดนี้ไปใช้ในทางสร้างสรรค์หรือไปปฏิรูปทฤษฎีวิวัฒนาการว่าต้องการจะใช้สิ่งที่มีชีวิตมีรูปแบบและคุณสมบัติเช่นใด

ย้อนหลังไปเมื่อ 20 ปีที่แล้ว เมื่อคำว่าการเชื่อมยีน (gene splicing) เริ่มเข้ามาเป็นคำศัพท์ใหม่ในเรื่องของมนุษยชาติ ก็ได้มีผู้ประนามว่าเทคโนโลยีเช่นนี้จะพาโลกไปสู่ความเสื่อมเพราะการพยายามเปลี่ยนแปรหน่วยของสิ่งมีชีวิตเช่นยีนโดยไม่คำนึงถึงความผิดชอบชั่วดีนั้นเป็นการผิดศีลธรรมอย่างร้ายแรง จนกระทั่งมีผู้นำทางศาสนาและหน่วยงานต่าง ๆ ของภาครัฐต้องยื่นมือเข้ามากำหนดแนวทางในเรื่องของการผิดกฎหมายเมื่อใครสักคนกระทำการปรับปรุงเพื่อให้ชีวิต (เช่น ยีน) ดีขึ้น อันเป็นกระทำที่ไม่มีใครรู้ว่าจะส่งผลแก่มวลมนุษย์อย่างไร หมายถึงว่าข้อมูลที่เป็น genetic information นั้นเป็นสิ่งที่ควบคุมยากเหมือนกับการควบคุมตรวจตราข้อมูลข่าวสารอื่น ๆ ซึ่งก็เป็นที่รู้กันว่าในอนาคตข้างหน้ามนุษย์จะสามารถกำหนดอนาคตของตนได้ด้วยการเลือกว่าจะให้ใครเกิดมาเป็นอะไร หรือมีรูปร่างหน้าตาตลอดจนมีความรู้สึกนึกคิดเช่นใด โดยอาศัยผลพวงของการค้นคว้าในยุคที่คอมพิวเตอร์เทคโนโลยีเฟื่องจนถึงจุดสุดยอด ดังกรณีของ Human Genome Project ที่เริ่มมาหลายปีแล้ว และได้มีการคาดกันว่าจะสามารถผลิตดอกออกผลไว้ในระหว่าง ค.ศ. 2001-2005 ซึ่งเมื่อถึงเวลานั้นก็ไม่มีใครรู้ว่ามนุษย์จะเป็นเช่นใด นอกจากจะมีมาตรการสำหรับใช้ในการบังคับและควบคุมในเรื่องนี้

โมเดลของชีวิตในยุคของซุปเปอร์คอมพิวเตอร์

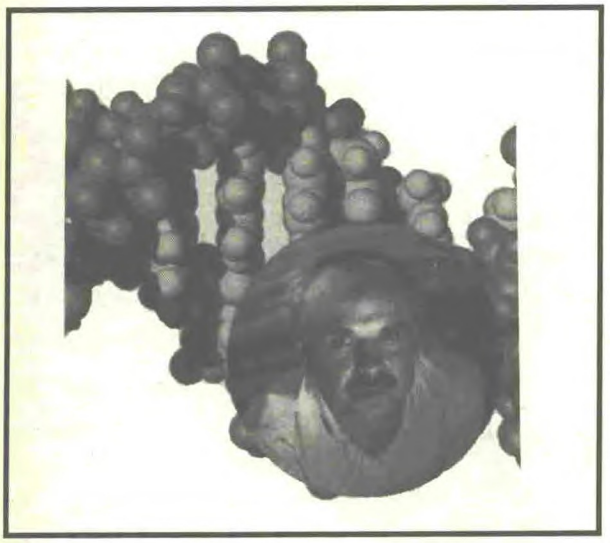
เมื่อประมาณ 40 ปีมาแล้ว เจมส์ วัตสัน (James Watson) กับ ฟรานซิส คริก (Francis Crick) เป็นบุคคลกลุ่มแรกที่สามารถหาโครงสร้างของ DNA ได้ด้วยการเกณฑ์คนให้มาช่วยกันประกอบชิ้นส่วนของโมเดล ซึ่งต้องใช้เวลานานนับเป็นแรมวันเพื่อพิสูจน์ว่าปรมาณูต่าง ๆ จับตัวกันตามที่นักวิทยาศาสตร์คิดว่าน่าจะเป็น จนกระทั่งถึงกลางทศวรรษ 1970 นักเคมีก็ยังอาศัยโมเดลทางกายภาพโดยการสร้างองค์ประกอบทางเคมีแบบที่ต่อกันเข้าทีละชิ้นแล้วใช้บรรทัดในการวัดเทียบโครงสร้างจำลองที่สร้างด้วยมือโดยเทียบกับเงาภาพอันไม่สู้จะชัดของโมเลกุลจริง ๆ ที่ได้จากเอ็กเรย์ (X-ray crystallography)

ต่อมาคอมพิวเตอร์ได้เข้ามาเปลี่ยนแปลงสิ่งที่พูดในข้างต้นทั้งหมดดังเช่นกรณีของมหาวิทยาลัยแห่งแคลิฟอร์เนียที่แซนดีโก้ (San Diego) ซึ่งอาศัยซุปเปอร์คอมพิวเตอร์ (เช่น Cray C90, Cray T3D รวมทั้ง Intel Paragon ที่มีโปรเซสเซอร์ 400 ตัว) เข้าช่วยในการค้นคว้าในเรื่องของพฤติภาพของ DNA โดยใช้ชื่อห้องทดลองว่า San Diego Supercomputer Center (SDSC) งานค้นคว้าที่สำคัญอันหนึ่งที่ SDSC ก็คือการศึกษาความเป็นไปของ **โมเลกุลของโปรตีน** ซึ่งต้องอาศัยการแสดงภาพ

เป็นแบบสามมิติเข้าช่วยอย่างมาก อันเนื่องจากว่าในความเป็นจริงนั้นโครงสร้างที่มีความซับซ้อนมากเช่นโปรตีนโมเลกุลไม่ได้หยุดนิ่งอยู่กับที่ แต่จะมีการเคลื่อนไหวอยู่ตลอดเวลาในรูปของ large-scale motions อันทำให้ต้องอาศัยเครื่องคำนวณในระดับซุปเปอร์คอมพิวเตอร์มาจำลอง (simulate) พฤติภาพเช่นนี้ คอมพิวเตอร์ที่มีความเร็วสูงจึงเข้ามาทำให้สิ่งที่ดูเหมือนจะอยู่นิ่ง ๆ เกิดมีชีวิตขึ้นมาในรูปของการเคลื่อนตัวอย่างช้า ๆ โดยอาศัยอุปกรณ์คำนวณ เช่น Paragon ที่ตั้งอยู่ในห้องทดลองที่สามารถให้ภาพที่เป็นแบบ 3-D บนสกรีนของ Silicon Graphics เวอร์กสเตชัน ภาพที่เห็นในจอแบบสามมิติจึงเต็มไปด้วยสีสันและมีชีวิต เหมือนก้อนวุ้นกลม ๆ เป็นจำนวนมากมายที่เป็นองค์ประกอบของโปรตีนโมเลกุลอันเป็นการเคลื่อนไหวที่มีการสั่นกระเพื่อมเป็นระยะ ๆ ปรากฏให้เห็น การที่ต้องมีการทุ่มเทอย่างมากในเรื่องนี้ก็เพราะว่า DNA ควบคุมการสร้างโปรตีนโมเลกุล และถ้านักวิจัยสามารถที่จะเห็นว่าองค์ประกอบขั้นพื้นฐานที่เป็นอนุภาคเล็ก ๆ นี้มีความสัมพันธ์กันอย่างไร ก็จะสามารถเข้าใจพฤติภาพของโมเลกุลของโปรตีนได้โดยไม่ยาก นอกจากนี้ยังสามารถใช้ข้อมูลที่ได้จากการทดลองในเรื่องนี้ไปใช้ในการพยากรณ์ว่าสารประกอบต่าง ๆ จะจับตัวกันจนเป็นมรรคผลได้อย่างไร อันทำให้สามารถสร้าง**ฐานข้อมูล** ออนไลน์ที่มีปริชาญาณได้ ซึ่งจะเป็นแหล่งข้อมูลขนาดใหญ่ที่นักวิทยาศาสตร์ที่ค้นคว้าในด้านนี้เข้าใจข้อมูลอันมีค่าในเรื่องของ DNA ได้โดยสะดวก ที่สำคัญก็คือว่าฐานข้อมูลที่ SDSC จะทำให้นักวิจัยสามารถแยกแยะและเปรียบเทียบโครงสร้างของโมเลกุลได้อย่างฉลาดโดยการใช้ข้อมูลที่มีอยู่แล้วให้เกิดประสิทธิผลในท่ามกลางปริมาณข้อมูลในด้านนี้ที่เพิ่มเป็นสองเท่าในทุก 18 เดือน ซึ่งก็เป็นสาเหตุอีกประการที่ทำให้ต้องอาศัยเครื่องคำนวณที่มีความเร็วสูงพร้อมด้วยซอฟต์แวร์ที่ใช้ในการสร้างโมเดลแบบสามมิติพร้อมด้วย search, algorithms ที่ผนวกเอาเทคนิคเช่น fuzzy logic ฯลฯ มาใช้สนับสนุนการค้นคว้านี้ จากนี้ก็จะเล่าถึงความเป็นมาก่อนการก้าวไปสู่ การสร้างแผนภูมิของ DNA

ในฤดูร้อนของปี 1980 ศูนย์อายุรศาสตร์ของ UCLA ได้ประกาศความสำเร็จการเปลี่ยนยีน (gene transplant) ของสัตว์ที่ประสบความสำเร็จเป็นครั้งแรก สิบปีต่อมานายแพทย์ที่ National Institutes of Health (NIH) ได้ทำการบำบัดโรคด้วยการเปลี่ยนยีนให้แก่มนุษย์เป็นผลสำเร็จเป็นครั้งแรก (first gene therapy) โดยใช้ DNA ที่สกัดมาจากมนุษย์ด้วยกันมาทำการปรับเปลี่ยนแล้วนำไปใส่ในตัวผู้ป่วย จากนั้นการทดลองในด้านนี้ก็มีมากขึ้นเรื่อย ๆ จนกระทั่งในเดือนกันยายน 1996 สำนักงานอาหารและยา (FDA) ของสหรัฐฯ ได้อนุมัติกรรมวิธีที่มีมากถึง 160 อย่างที่ใช้ในการรักษาโรคที่ร้ายแรงเช่นมะเร็งที่ปอด จนถึงโรคโลหิตจาง นอกจากนี้การค้นในแนวนี้ยังได้แพร่คลุมไปถึงด้านการเกษตร เช่น การ antifreeze gene ของปลาที่อยู่บริเวณอาร์คติก ไปใส่ในพืชถั่วเพื่อให้พืชชนิดนี้ทนความหนาวเย็นได้ในประเทศหนาว อันเป็นการปูทางไปสู่การค้นคว้าในเรื่องของ mapping DNA ซึ่งก็ต้องขอเล่าย้อนกลับไปเมื่อราว 10 ปีที่แล้วอันเป็นช่วงแรก ๆ ของการค้นคว้าเพื่อหามรรคผลจาก DNA ให้มากขึ้น อันเป็นธุรกิจที่เต็มไปด้วยความสับสนโดยไม่มีแนวทางที่แน่นอนแต่อย่างใด

ในปี 1985 ได้มีการประชุมทางวิชาการที่มหาวิทยาลัยแห่งแคลิฟอร์เนียที่ซานตาครูส (Santa Cruz) เพื่อวางแผนรวมในอันที่จะหาความกระจ่างในเรื่องของ sequence ของอนุภาคของชีวิตที่จะนำไปสู่การค้นพบ **เจนโนมของมนุษย์** (entire human genome) ซึ่งผู้อ่านคงไม่ลืมว่า คำว่า genome นั้นหมายถึง **"เจเนติกโค้ดที่สมบูรณ์แบบที่ใช้ในการสร้างมนุษย์ขึ้นมา"** จากนั้นอีกหนึ่งปีกระทรวงพลังงานของสหรัฐฯ ก็ได้แสดงความสนใจในเรื่องนี้อันเป็นผลเนื่องมาจากการศึกษาวิเคราะห์ในด้านความเสียหายของยีน (genetic damage) เนื่องมาจากการแผ่รังสีนิวเคลียร์ จากนั้นอีกหนึ่งปี NIH ได้เข้ามาเกี่ยวข้องในเรื่องนี้พร้อมด้วยหน่วยงานอื่นในสหรัฐฯ อีกสองหน่วยงานที่ร่วมกันประกาศเพื่อก่อตั้งโครงการที่เรียกว่า Human Genome Project อันเป็น **สหกิจกรรมขนาดมหึมา**ที่เริ่มเมื่อปี 1990 และไปสิ้นสุดเมื่อปี 2005 อันเป็นการค้นคว้าที่ต้องร่วมมือกันแบบข้ามชาติที่มีประเทศญี่ปุ่นรวมอยู่ด้วย



รูปที่ 1 Mark Boguski หนึ่งในบรรดา computational biologists ที่ร่วมโครงการ human genome ที่มีความเห็นว่า ใน ค.ศ. 2003 หรือ 2004 ค่าใช้จ่ายในการจัดระเบียบ (sequencing) DNA จะลดลงจนทำให้การนำคอมพิวเตอร์ขนาดเล็กเข้ามาประมวลผลเพื่อสร้างมนุษย์ในอุดมการณ์เป็นไปได้สำหรับคนส่วนมากที่มีคอมพิวเตอร์ในครอบครอง

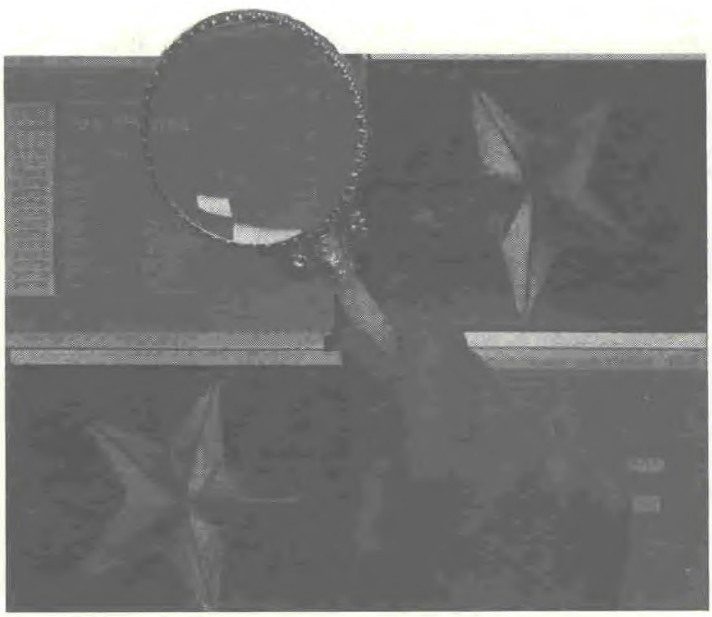
นอกเหนือจากการขอความร่วมมือจากห้องทดลองต่าง ๆ ที่กระจายอยู่ทั่วโลก ปัญหาในเรื่องของการค้นคว้าที่เกี่ยวกับ DNA ก็คือว่าเป็นการค้นคว้าทดลองที่แสนแพงอันเป็นการแย่งงบประมาณไปจากการค้นคว้าในด้านอื่นที่ดูเหมือนว่ามีความสำคัญกว่า ที่แน่ ๆ ก็คือว่าการค้นคว้าในด้าน DNA นั้นได้มีข้อมูลที่เข้ามาเกี่ยวข้องด้วยในปริมาณอันมหาศาลจนอุปกรณ์ที่ใช้ในการจัดเรียงอนุภาคในระดับโมเลกุล (sequencers) ต้องทำงานอย่างหนักและทรุดโทรมในเวลาอันสั้น

ตัวอย่างของฐานข้อมูลที่ตั้งอยู่ที่ SDSC ก็มี เช่น การเก็บข้อมูลด้านโมเลกุลที่เป็นแค่เพียง **one family of molecules** ซึ่งมีนักวิทยาศาสตร์ผู้หนึ่งเป็นผู้ดูแลอยู่ อันเป็นข้อมูลที่ใช้ในการค้นคว้าด้วยคอมพิวเตอร์ที่เป็นของบุคคลเพียงคนเดียว แต่ในทางรวม ๆ นั้นการค้นคว้าที่ทำกันเป็นทีมทั้งในและนอกประเทศเช่นในกรณีนี้จะต้องการห้องสมุดอิเล็กทรอนิกส์ที่ใหญ่มาก ๆ เพื่อให้สามารถเก็บข้อมูลทั้งหมดจาก Human Genome Project รวมทั้งข้อมูลอื่นที่เป็น DNA sequences จาก species นับเป็นพันที่เป็นของพืชและสัตว์ ทำให้ต้องมีแหล่งข้อมูลขนาดยักษ์อยู่สามแห่ง แห่งหนึ่งอยู่ในโตเกียว อีกแห่งอยู่ที่เคมบริดจ์ในประเทศอังกฤษ และอีกแห่งอยู่ที่ GenBank ในเบเธสดะ (Bethesda) แมรี่แลนด์ในสหรัฐฯ ซึ่งอยู่ภายใต้การดูแลของ the National Center for Biotechnology Information (NCBI) ที่ได้รับเงินอุดหนุนจาก NIH ฐานข้อมูลต่าง ๆ ของทั้งสามแหล่งนี้ต่างแลกเปลี่ยนและแบ่งกันใช้ข้อมูลอยู่ตลอดเวลา ทำให้ข้อมูลในทุก ๆ ด้านมีความสมบูรณ์ในทุกขณะ ความโดดเด่นของ GenBank นั้นนอกจากจะเก่าแก่ที่สุดแล้ว ยังได้ชื่อว่าเป็นแหล่งเก็บข้อมูลด้าน genetic ที่มีความเป็นต่อในแง่ของการนำข้อมูลดังกล่าวไปใช้ประโยชน์โดยอาศัยความชำชองแบบสุด ๆ และอาศัยคอมพิวเตอร์ที่มีอานุภาพเข้าช่วย ในกลางปี 1996 คอมพิวเตอร์ของ GenBank ได้เก็บข้อมูลในด้านยีนไว้ 700 ล้านคู่จาก species ต่าง ๆ นับเป็นจำนวน 1800 species การเก็บข้อมูลเช่นนี้ทำได้ไม่ยาก ความยากที่นักวิทยาศาสตร์และผู้ค้นคว้าต่าง ๆ เผชิญก็คือการตีความในข้อมูลดังกล่าว

GenBank อันเป็นสุดยอดของฐานข้อมูลด้าน DNA

อันดับอันเป็นแถวของ DNA ที่อยู่ในลักษณะโดด ๆ นั้นจะไม่มีความหมาย หากไม่รู้ว่า DNA sequence นั้นมาจากไหน และมีหน้าที่อะไร ซึ่งถ้าจะมองเรื่องนี้ในทางประวัติศาสตร์ก็จะเห็นว่า National Library of Medicine ของสหรัฐฯ ได้เริ่มเก็บข้อมูลทางการแพทย์และอื่น ๆ ที่เกี่ยวกับโรคภัยมาตั้งแต่สมัยสงครามกลางเมือง โดยเริ่มจากการบันทึกไว้ในเอกสารที่แพทย์ใหญ่ของกองทัพบกเป็นผู้ดูแล จนในปัจจุบันการเก็บข้อมูลเช่นนี้ได้มีดัชนี (indexes) ที่เกี่ยวโยงกับเรื่องของโรคภัยมากถึงปีละ 400,000 เรื่อง ว่าได้มีการค้นพบอะไรบ้างในทางการแพทย์ การเก็บบันทึกข้อมูลตั้งแต่อดีตจนปัจจุบันอย่างต่อเนื่องในรูปของ sequence ต่าง ๆ นี้ได้หมายถึงว่า หน้าที่ของฐานข้อมูลนั้นไม่เพียงแต่เก็บบันทึกเหตุการณ์ความเป็นมาเช่นในทางการแพทย์ให้อยู่ในรูปของ sequence เท่านั้น แต่ยังเชื่อมโยงสิ่งที่เก็บไว้ให้เข้ากับการค้นพบใหม่ ๆ ที่ทำให้ฐานข้อมูลมีความทันสมัยหรือมีการเติบโตฉลาดเฉลียว อันทำให้นักวิจัยในปัจจุบันสามารถใช้ประโยชน์ในการค้นพบสิ่งใหม่ ๆ ได้ ดังเช่นกรณีของนักวิทยาศาสตร์ที่ NCBI ที่ดูแล GenBank อยู่ เช่น นักวิทยาศาสตร์ผู้มีนามว่า **Mark Boguski** ซึ่งเป็นผู้ให้อรรถาธิบายในตอนแรกของหัวข้อนี้ว่า sequence ของ DNA ที่อยู่อย่างอิสระนั้นไม่มีความหมายโดย Boguski ได้อธิบายต่อไปว่าสมมติว่า ใครคนหนึ่งสามารถสร้างยีนสำหรับโรค Algheimer และต้องการจะดูว่าที่ GenBank มี sequences อะไรบ้างที่มีลักษณะเหมือนกัน เขาอาจพบ sequences เช่นนี้ในหนอน nematode ซึ่งทำให้เขาสามารถทำการทดลองกับ nematodes จนสามารถพบที่ใช้ในสมองมนุษย์ ตัวอย่างอีกอันอาจเป็นว่านักวิทยาศาสตร์ต้องการจะเรียนรู้ยีนในตัวของหนูเป็นการเพิ่มเติม และจากการค้นหาในฐานข้อมูลเช่น GenBank ก็พบว่ามีโปรตีนที่เหมือน ๆ กันที่มีอยู่ในยีสต์ ทำให้เขาสามารถทำการทดลองต่อไปโดยอาศัยกลุ่มของยีสต์แทนที่ต้องทดลองกับหนูอันทำให้ลดค่าใช้จ่ายลงได้มาก จนในที่สุดผลการทดลองก็จะทำให้เขาสามารถนำผลที่ได้ไปเทียบกับยีนของมนุษย์ที่มีความคล้ายคลึงกันเป็นวัฏจักรครบวงจร

คำว่า "เหมือน" หรือ "คล้ายคลึงกัน" (similar) ดังในข้างต้นดูเป็นสิ่งที่ง่าย แต่ความจริงแล้วเป็นของยาก กุญแจในที่นี้จึงอยู่ที่ความเหมือนหรือความคล้ายคลึง ยีนตัวหนึ่งอาจทำหน้าที่เดียวกันใน species ต่าง ๆ แต่จะไม่มีโครงสร้างที่เหมือนกัน เพราะในกระแสแห่งวิวัฒนาการนั้นจะเต็มไปด้วยการทดลองถูกและทดลองผิด sequence ของ DNA ที่อยู่ในกระแสของวิวัฒนาการที่ใช้เวลานับล้านปีที่แปรสภาพไปเรื่อย ๆ เพื่อให้การรับรองว่าทุกสิ่งย่อมมีความแตกต่างกันอันเป็นธรรมดา ปัญหาจึงมีว่าเราจะรู้ได้อย่างไรว่ายีนสองตัวนั้นสามารถเปรียบเทียบกันได้ ซึ่งนักวิทยาศาสตร์ที่ GenBank ก็แก้ปัญหาได้โดยอาศัยอัลกอริธึมอันชาญฉลาดเพื่อหาความคล้ายคลึงของยีนสองตัว แต่ในขั้นสุดท้ายก็ต้องทำการทดลองเพื่อยืนยันข้อเท็จจริงในห้องทดลองอีกที โดยสมมติว่าท่านเชื่อว่ายีนของคนกับยีนของยีสต์ (yeast) ทำหน้าที่เดียวกัน ก็ต้องมีการพิสูจน์โดยสกัดยีนออกมาจาก DNA ของยีสต์แล้วนำไปแทนยีนของมนุษย์ จากนั้นก็รอดูว่ายีสต์ยังเติบโตตามปกติหรือเปล่า อันเป็นกรรมวิธีที่เคยประสบความสำเร็จในการทดลองกับยีนชนิดต่าง ๆ มาราว 75 ครั้ง อันเป็นอรรถาธิบายของ Boguski ที่ได้แถลงต่อไปว่า นักชีววิทยาได้เคยชินกับกรรมวิธีที่เรียกว่า wait-and-see มานานแล้ว อันเป็นวิธีการที่บรรดาคอมพิวเตอร์โปรแกรมเมอร์ไม่เห็นด้วย เพราะขาดความแม่นยำ เพราะกระบวนการทางธรรมชาตินั้นจะซับซ้อนเกินกว่าที่ใครจะคาดถึง จนทำให้เกิดมีความต้องการสาขาอาชีพที่เรียกว่า computation biologists ขึ้นมา ที่ทำให้นักคอมพิวเตอร์ทั้งหลายตั้งวิ่งเข้าไปในสาขานี้ แต่ตามความเห็นของนักค้นคว้าชั้นนำ เช่น Boguski ได้เห็นว่า สิ่งที่ต้องการมาก ๆ ในขณะนี้ก็คืออันดับแรกต้องเป็นนักชีววิทยาและตามด้วยความเป็นนักวิทยาศาสตร์ด้านคอมพิวเตอร์ ส่วนตัวของ Boguski นั้นก็ได้เริ่มด้วยปริญญา MD เป็นอันดับแรก แล้วหันมาเป็นนักคอมพิวเตอร์ในภายหลัง อันจำเป็นต้องเลือกเอาในระหว่างการเอาใจผู้ป่วยกับอุดมการณ์ที่ว่าจะเลือกเดินทางไหนจึงจะก่อให้เกิดผลกระทบกับคนส่วนใหญ่ให้มากที่สุด ซึ่งในที่สุดเขาก็เลือกที่จะเดินไปในดงแดนของ DNA sequences ที่เขาเห็นว่าจะเป็น **อายุรศาสตร์ในอนาคต** อันเป็นที่มาของ**หอสมุดทางการแพทย์**ในศตวรรษที่ 21 ดังจะเล่าต่อไป



แนวคิดสำหรับฐานข้อมูลยุคใหม่

สถาปนิกเอกของหอสมุดดังกล่าวในข้างต้นก็คือจิม ออสเตล (Jim Ostell) ซึ่งเป็นผู้อำนวยการการออกแบบขั้นพื้นฐานของฐานข้อมูลที่ GenBank ในปี 1990 ความรู้หลักของออสเตลคือสัตว์วิทยา (zoology) ในขั้นปริญญาตรี จากนั้นก็ละทิ้งโลกของวิชาการโดยหันมามีอาชีพในการกว้านหาสิ่งมีชีวิตทั้งพืชและสัตว์มาขายให้ supply house ในปี 1972 ทำให้เขาต้องท่องไปในที่ต่าง ๆ ในมลรัฐแมสซาจูเซตต์ที่พอจะเก็บเกี่ยวสิ่งมีชีวิตทั้งพืชและสัตว์ไปขายให้ผู้ซื้อที่เรียกว่า biological supply house จากนั้นก็หันหลังให้อาชีพนี้เพื่อไปทำปริญญาโท เนื่องจากได้เห็นความซับซ้อนของสิ่งมีชีวิตที่เกิดจากจุลชีพที่มีรูปแบบอันแสนง่าย ในปี 1979 ออสเตลเริ่มทำงานในห้องทดลองทางชีววิทยาของฮาร์วาร์ด ซึ่งในขณะนั้นก็ได้มีการเริ่มงานในด้าน DNA sequencing กันบ้างแล้ว โดยคอมพิวเตอร์ที่มีใช้ในห้องทดลองของฮาร์วาร์ดในขณะนั้นก็เป็นแค่เพียง CP/M แมชีนที่เทียบได้กับ Radio Shack TRS-80 ที่เลขาฯใช้ทำงานด้านเวิร์ดโปรเซสซิงในตอนกลางวัน ส่วนโปรแกรมที่มีอยู่เพื่อใช้ในการจัดการกับข้อมูลทางเคมีก็แสนจะปัญญาอ่อน ทำให้ออสเตลต้องเขียนโปรแกรมขึ้นมาเองพร้อมทั้งแจกจ่ายให้ผู้สนใจ ซึ่งเมื่อหลังจากนี้เขาได้ Ph.D. ก็อำลาจากห้องทดลองเพื่อไปขายซอฟต์แวร์ที่เขาพัฒนาขึ้นมา ในช่วงขณะนั้นออสเตลได้ใช้ชีวิตอยู่ในบ้านนาที่เวอร์มอนต์ (Vermont) อย่างสงบเพื่ออาศัยความสันโดดในการสร้างสรรค์สิ่งที่เป็นคอมพิวเตอร์โค้ดต่อไป แต่อยู่ในฟาร์มได้ไม่นาน ออสเตลก็มาทำงานกับ NCBI ในปี 1988 ในสาขา molecular biology ในตำแหน่ง Senior Biomedical Research Scientist โดยมาอยู่ที่ GenBank ซึ่งเป็นที่เก็บข้อมูลในด้านชีววิทยาระดับโมเลกุลอยู่ 3 ลักษณะคือ

- DNA sequences
- Protein sequences ที่เกิดจากยีน
- เอกสารทางวิชาการด้านพันธุกรรม (genetic) และอื่น ๆ ที่เกี่ยวข้องจากฐานข้อมูล Medline ของ NIH อันเป็นแหล่งข้อมูลสำหรับงานวิจัยทางการแพทย์แบบออนไลน์ที่สมบูรณ์ที่สุดในโลก

สรุปก็คือว่า GenBank มีทุกอย่างตามที่กล่าวในข้างต้น ไม่ว่าผู้ค้นคว้าจะเริ่มจากจุดไหน เพราะทุกสิ่งที่เก็บอยู่ในฐานข้อมูลจะเชื่อมโยงกันหมด แม้กระทั่งเอกสารทางวิชาการเก่า ๆ ก็ยังเชื่อมโยงจากอดีตมาสู่ปัจจุบัน โดยมาเชื่อมโยงกับเอกสารทางวิชาการใหม่ ๆ อานุภาพของฐานข้อมูลที่เก็บอยู่ที่ GenBank นี้จะเป็นในทำนองที่ว่ามักจะให้คำตอบหรือข้อมูลที่เป็นประโยชน์แก่ผู้ใช้โดยใกล้เคียงกับความต้องการที่สุดในวาระแรกที่เรียกใช้ข้อมูลในเรื่องใดเรื่องหนึ่ง (displays the most promising hits first)

หลักการทำงานของฐานข้อมูลที่นายออสเตลรับผิดชอบอยู่ก็คือ จะมีการคำนวณเพื่อเปรียบเทียบโปรตีนทั้งหมดซึ่งเป็นที่รู้จักกันดี กับโปรตีนอื่นที่รู้จัก โดยอาศัยหลักการอันเดียวกับการเปรียบเทียบตัวหนังสือหรือ text คือในเอกสารแต่ละชิ้นจะมีการให้น้ำหนักสำหรับคำต่าง ๆ ซึ่งขึ้นกับว่าคำ ๆ นั้นจะเกิดซ้ำกันบ่อยแค่ไหน โดยได้ดำเนินการในทำนองเดียวกันนี้กับ คำย่อ (citations) นับเป็นจำนวน 1.2 ล้านคำย่อของเอกสารทางวิชาการที่เป็นส่วนหนึ่งของเรื่องพันธุกรรมที่เก็บไว้ใน Medline ข้อมูลอื่นที่จะต้องดำเนินการในแบบเดียวกันนี้ก็คือข้อมูลด้านพันธุกรรม (outside genetics) อีก 8 ล้านชิ้นที่ได้มาจากที่อื่นที่อยู่นอก NIH ประโยชน์ที่ได้ก็คือ จะทำให้นักวิจัยสามารถนำ DNA sequence ที่พบไป**เปรียบเทียบ**กับ DNA sequence อื่นที่อาจรู้จักดีกว่า การเชื่อมโยงเช่นนี้ยังสามารถนำไปสู่อาการของโรคที่เกิดจากในการที่ sequence อันใดอันหนึ่งเกิดหายไปหรือเกิดความเสียหาย จนในขั้นสุดท้ายก็อาจเชื่อมโยงไปถึงการบำบัดต่าง ๆ ที่จะนำมาปรับปรุงและใช้รักษาโรคให้หาย

ระบบของ GenBank จะอาศัยคอมพิวเตอร์ที่เป็นแบบ Sun SPARC center 2000 ซึ่งมีอยู่หลายเครื่อง รวมทั้งเครื่องแบบ Power Challenges XL อีก 4 ตัวที่ทำหน้าที่เหมือน Sun ในเรื่องของ sequence searches นอกจากนี้ก็มีพนักงานประจำอยู่ 3 คนที่ให้บริการแก่ผู้ใช้โดยตอบคำถามทาง e-mail โทรศัพท์ หรือแฟกซ์ การรับข้อมูลของ GenBank ได้อาศัยโมเด็มที่ทำให้สามารถรับข้อมูลที่เป็น DNA sequences ที่ได้มีการทำความเข้าใจแล้วโดยมาจากแหล่งต่าง ๆ นับเป็นร้อยแห่งที่

กระจายอยู่ทั่วโลก มีอยู่บ่อยครั้งที่ห้องทดลองต่าง ๆ มักจะศึกษายีนตัวเดียวกัน เว้นแต่ว่ายีนดังกล่าวจะไม่เหมือนกันยังกับแกะ อันเนื่องมาจากข้อแตกต่างในด้าน genetic เพียงเล็กน้อยในระหว่างการเก็บตัวอย่างของยีนแต่ละตัว ทำให้ GenBank ต้องเก็บข้อมูลของเวอร์ชันต่าง ๆ ที่เป็นของ sequence ต่าง ๆ โดยไม่มีคู่ใดเหมือนกับแบบ 100 เปอร์เซ็นต์

การตีความในยุคที่มีการชุบชีวิต

เมื่อ Human Genome Project สิ้นสุดลงก็จะได้ sequence สุดท้ายที่คัดมาจากชิ้นส่วนย่อย ๆ นับเป็นพันชิ้นในลักษณะของการผนวกผสมแบบหลายเผ่าพันธุ์ สิ่งที่ได้ดำเนินการมาถึงปัจจุบันก็ถือการ sequence เจนโนมของมนุษย์ได้เพียงร้อยละ 5 ของทั้งหมดที่ต้องทำการ sequence อันเปรียบได้ว่าชีววิทยาในปัจจุบันก็เหมือนกับวิชาฟิสิกส์ก่อนนิวตันในแง่ที่ว่า นักวิทยาศาสตร์เริ่มมองเห็นแพตเทอร์นที่มีความหมายในข้อมูลที่เป็นกองภูเขา และเมื่อมนุษย์สามารถเข้าใจ DNA ได้ก็จะไปสู่ยุคของการปฏิรูปทางอิเล็กทรอนิกส์ อันหมายถึงว่า species ต่าง ๆ จะสูญสิ้นชีวิตในขั้นนี้เพราะมีพละอำนาจพอที่จะกำจัดตัวเองได้ ได้มีผู้ให้ความเห็นต่อไปอีกว่าธรรมชาติได้สร้างสรรพสัตว์ที่ดูน่ากลัวไว้มากมายกว่าที่มนุษย์จะนึกถึง จนมนุษย์ต้องตกเป็นเหยื่อที่มีข้อเสียเปรียบกว่าสรรพสัตว์อีกหลายชนิด ในบรรดาศัตรูที่น่าเกรงขามที่สุดของมนุษย์นั้นน่าจะเป็นไวรัสที่ถือว่าเป็น**นักล่า** อันทำให้มนุษย์ต้องทำความเข้าใจนักล่าให้มากที่สุด เพื่อจะได้ป้องกันตนเองได้ ปัญหาจึงมีว่าควรจะเปิดเผยข้อมูลเท่าที่พูดมานี้มากแค่ไหนในแง่ของสาธารณะ กล่าวอีกนัยหนึ่งก็คือ**ใครเป็นเจ้าของ DNA** การตอบคำถามนี้เป็นเรื่องใหญ่ซึ่งต้องพาดพิงไปถึงกฎหมาย เช่น The Genetic Privacy Act ของสหรัฐฯ หรือพาดพิงไปถึงองค์กร เช่น DNA Advisory Committee ซึ่ง NIH เป็นผู้ประกาศตั้ง รวมทั้งองค์กรที่เรียกว่า The National Center for Human Genome Research และหน่วยงานของ Unesco เช่น International Bioethics Committe ที่มีกำเนิดขึ้นมาในเดือนมีนาคม 1996 และได้มีประกาศที่เรียกว่า **"Declaration on Protection of the Human Genome"** โดยมีแนวนโยบายว่างานค้นคว้าทั้งหมดที่เกี่ยวกับ genetic ควรมีรัฐบาลเป็นผู้กำกับดูแล อันเนื่องจากว่า **เจนโนมของมนุษย์เป็นมรดกร่วมกันของมนุษย์ชาติ** นอกจากนี้ก็ยังมีความแตกต่างอย่างอุกฤติสำคัญ

ระหว่างคำว่า **การแก้ไข** (correction) กับ**การปรับเสริมให้ดีขึ้น** (enhancement) ในแง่ของการตีความในทางปฏิบัติ เพราะอาจมีการถกเถียงโต้แย้งได้มากในยุคทองของ DNA เพราะชีวิตของมนุษย์อันสุดแสนจะซับซ้อนนั่นเอง ดังตัวอย่างต่อไปนี้

ในกรณีของเด็กที่มีแนวโน้มทางกรรมพันธุ์ว่ามีโอกาสจะอ้วนและอายุสั้นกว่าคนส่วนใหญ่ที่ไม่อ้วน จึงต้องมีการตีความว่า**ความอ้วน**เป็นโรคอย่างหนึ่ง (หรือไม่) ที่ต้องทำการแก้ไข (correct) ใน germ cells หรือเป็นการกระทำที่อยู่ในรูปของการปรับเสริมให้ดีขึ้น (enhancement) รวมไปถึงเรื่องสายตาสั้นหรือยาวที่มีผลต่อการประกอบอาชีพ เช่นการเป็นนักบิน ตลอดจนความสูงหรือความเตี้ย ที่อาจมีผลในเรื่องของการปฏิบัติงาน เส้นกั้นระหว่างคำว่า enhanced และ corrected จึงไม่เด่นชัด ซึ่งในบางกรณีก็อาจเลือนลางอย่างมากจนยากแก่การตีความ โดยในบางครั้งก็อาจมีเรื่องของศีลธรรมเข้ามาเกี่ยวข้องด้วย ซึ่งล้วนเป็นปัญหาสำหรับในอนาคตอันใกล้ทั้งสิ้น เมื่อถึงตอนนั้นก็อาจต้องอาศัยบุคคลในหลายสาขาวิชาชีพมานั่งถกเถียงหรือตีความว่าอะไรคือ correction อะไรคือ enhancement ซึ่งก็จะเหมือนกับปัญหาสังคมอื่น ๆ ที่มีอยู่ในขณะนี้ที่ยากจะวินิจฉัยให้แน่ชัดลงไปว่าอะไรถูก อะไรผิด นอกเสียจากจะเป็นการวินิจฉัยแบบยกประโยชน์ให้ผู้ที่กำลังประสบเคราะห์กรรมหรือต้องการมีชีวิตที่ดีขึ้น โดยใช้วิธีการรอมชอมให้มากที่สุดเท่าที่จะทำได้ โดยอาศัยศีลธรรมเป็นพื้นฐาน

จากนี้ก็จะพูดถึงเรื่องของค่าใช้จ่ายในเรื่องของการ sequence DNA ซึ่งถือว่ามีผลกับการรุดหน้าในการค้นคว้าในเรื่อง DNA อย่างมาก

ค่าใช้จ่ายในการค้นคว้า

ผู้เชี่ยวชาญส่วนใหญ่ได้เห็นว่าแม้ว่าค่าใช้จ่ายในการ sequence DNA จะลดลงเรื่อย ๆ อันทำให้โครงการ เช่น Human Genome Project สามารถบรรลุผลก่อนหมายกำหนดและมีการใช้จ่ายที่อยู่ในขอบเขตของงบประมาณก็ตาม ปัญหาก็คือว่า เราจะไปถึงจุดที่การ sequence DNA จะมีราคาถูกลงจนกระทั่งผู้บริโภคหรือผู้ใช้สามารถที่จะนำ**เจนโนม**ของตนมารับการสแกนและเก็บผลเอาไว้ใน CD-ROM นั้นเป็นสิ่งที่เป็นไปได้หรือไม่ ซึ่งยังตอบยาก เพราะเป็นเทคโนโลยีที่ใหม่มาก ๆ การพยากรณ์เพื่อทราบ**ค่าใช้จ่าย**ในด้าน DNA sequencing จากในปัจจุบันก็เหมือนกับการพยายามคาดคะเนรูปแบบของคอมพิวเตอร์ในอนาคต โดยในขณะที่ทำการคาดคะเนนั้นเป็นทศวรรษ 1950 อันเนื่องจากว่าค่าใช้จ่ายนี้จะมีผลกับความรุดหน้าในเรื่องการใช้ประโยชน์จาก DNA

ผู้ที่จะสามารถให้ความกระจ่างในเรื่องของค่าใช้จ่ายในการค้นคว้าในเรื่องของการค้นคว้าในด้าน DNA ในส่วนที่ว่าด้วย large-scale mapping and sequencing ก็คือ เจน ปีเตอร์สัน (jane Peterson) ซึ่งเป็นผู้อำนวยการในด้าน genomic sequencing โดยประจำอยู่ที่ The National Center for Human Genome Research ที่มีหน้าที่จัดสรรเงินปีละ 119 ล้านเหรียญสหรัฐฯ ต่อปี แก่ห้องทดลองต่าง ๆ ที่เกี่ยวข้องกับ Human Genome Project เจนได้ให้ความเห็นว่าเมื่อปี 1990 อันเป็นปีที่โครงการนี้เริ่ม ค่าใช้จ่ายในการค้นคว้าในเรื่องนี้จะตกราว **10 เหรียญสหรัฐฯ ต่อ base pair** อันเป็นค่าใช้จ่ายที่แสนแพงเมื่อเทียบกับจำนวนเจนโนมในตัวคนที่มีมากถึง **3 พันล้าน base pairs** ต่อคนหนึ่งคน

หน้าที่อันอุกฤตสำคัญของเจน ปีเตอร์สันก็คือการหาทางลดค่าใช้จ่ายในการค้นคว้า ซึ่งก็ได้ลดลงเรื่อย ๆ อันเนื่องจากพัฒนาการทางเทคโนโลยีและประสิทธิผลอันเนื่องจากขนาดของงานค้นคว้าที่ใหญ่พอสมควร (efficiencies of scale) อันทำให้ค่าใช้จ่ายลดลงเหลือ 50 เซ็นต์ต่อ base pair สำหรับในปัจจุบัน ขณะนี้เจนได้แจกงบประมาณราว 42 ล้านเหรียญสำหรับโครงการขั้นบุกเบิกเป็นจำนวน 6 โครงการ เพื่อหากลยุทธ์และเทคโนโลยีใหม่ ๆ ที่จะทำให้ค่าใช้จ่ายในการค้นคว้าด้าน DNA ลดลงไปอีกจนถึง base pair ละ 10 เซ็นต์ อย่างไรก็ตาม จวบจนกระทั่งในขณะนี้งานด้าน mapping และ sequencing สำหรับ DNA เสร็จไปเพียงร้อยละ 5 เท่านั้น ค่าใช้จ่ายทั้งหมดสำหรับ complete sequence จะตกราว 300 ล้านเหรียญ และเมื่อถึงจุดที่ Genome Project ใกล้จะสิ้นสุดคือในราวปี 2003 หรือ 2004 ค่าใช้จ่ายในด้าน sequencing ก็จะตกราวไม่ถึงหนึ่งเซ็นต์ต่อหนึ่ง base pair ถ้าเป็นจริง งานด้าน sequencing ก็จะก้าวมาถึงจุดที่สามารถดำเนินการให้เฉพาะกับแต่ละบุคคลได้ในอีกราว 20-30 ปีข้างหน้า

ปัญหาก็คือว่าทำไมคนทั้งหลายจึงต้องการ**สำเนา** (copy) **ของข้อมูล DNA** ที่เป็นคุณสมบัติประจำตัวไว้ในครอบครองในรูปของฉบับของใครฉบับของมัน คำตอบก็คือว่า DNA นั้นไม่ได้เป็นเพียงแค่ข้อมูล แต่เป็นโปรแกรมอย่างหนึ่ง คือเมื่อเราสามารถอ่านและเข้าใจโปรแกรมนี้ได้ ผู้ที่มีคอมพิวเตอร์ทั้งหลายก็จะทำในสิ่งที่พึงกระทำเมื่อมีโปรแกรมอยู่ในมือ นั่นก็คือโหลดโปรแกรมเช่นนี้เข้าไปในคอมพิวเตอร์และทดลองวิ่งดู ที่สำคัญมาก ๆ ก็คือว่าการค้นคว้าวิจัยเท่าที่กล่าวมานี้จะไปเกี่ยวข้องกับการบำบัดโรคที่เรียกว่า germline therapy อันเป็นเรื่องของการเอาชนะความผิดปกติในร่างกายของมนุษย์ที่ต้องการการลงทุนอันแสนแพงดังกล่าวในข้างต้น

เรียกวิญญาณของไอน์สไตน์ให้คืนมา

ในอนาคต คอมพิวเตอร์โมเดลจะมีความซับซ้อนมากขึ้นเรื่อย ๆ อันไม่ได้หมายถึงเฉพาะแค่โมเดลของโมเลกุลเท่านั้น แต่ยังคลุมถึงโมเดลที่มีรายละเอียดอันเกี่ยวกับการทำงานของภายในเซลล์ การสร้างโมเดลเช่นนี้จึงเป็นงานท้าทายที่ถือว่ามีความยิ่งใหญ่เอามาก ๆ เพราะยังมีหน้าที่อีกหลาย ๆ อย่างภายในเซลล์ที่มนุษย์ยังไม่เข้าใจ อย่างไรก็ตาม นักวิทยาศาสตร์ยังอาจไม่ต้องการโมเดลของเซลล์ที่เป็นฉบับสมบูรณ์ในการทำเทียม (simulate)

สำหรับกระบวนการต่าง ๆ ที่เกิดขึ้นในเซลล์ กล่าวอีกนัยหนึ่งก็คือว่า นักวิทยาศาสตร์ควรมีความรู้พอที่จะสร้างโมเดลที่สามารถจำลองกระบวนการต่าง ๆ ภายในเซลล์ที่เห็นว่าเป็นไปได้ในทางชีววิทยาโดยไม่จำเป็นต้องเจาะลึกลงไปถึงพฤติภาพของแต่ละโมเลกุล อันจะเหมือนกับการที่นักอุตุนิยมวิทยาสามารถพยากรณ์ลักษณะอากาศโดยไม่ต้องมองไปถึงหยดน้ำฝน และเมื่อใดก็ตามที่นักวิทยาศาสตร์ด้านชีวภาพสามารถมีความเข้าใจดังเช่นที่ว่านี้ การค้นคว้าเบื้องต้นก็จะไม่ต้องอาศัยการทดสอบในห้องทดลองที่ต้องใช้เวลานานและสิ้นค่าใช้จ่ายสูง แต่จะสามารถใช้คอมพิวเตอร์เมมโมรีเข้าช่วยในการทดลอง โดยอาศัยกฎต่าง ๆ ที่แสดงการทำปฏิกิริยาทางเคมีหรือปฏิกิริยาอื่น ๆ เป็นปัจจัยในการทดลอง ผนวกกับการใช้คอมพิวเตอร์เทคโนโลยีในด้าน parallel processing ซึ่งให้สามารถติดตามความเป็นไปของกระบวนการต่าง ๆ ทางชีวเคมีอย่างพร้อม ๆ กันได้ในรูปของการแบ่งงานกันในการคำนวณ อันทำให้การซิมูเลตพฤติภาพของเซลล์หลาย ๆ เซลล์ในขณะแบ่งเซลล์และทำปฏิกิริยา ซึ่งจะนำไปสู่การปลูกสร้างจุลชีพที่มีชีวิตเป็นสิ่งที่เป็นไปได้ ทั้งหมดนี้ก็เป็นการพูดถึงการใช้อานุภาพของ parallel computer มาใช้ในการซิมูเลต การสร้างชีวิตด้วยเซลล์ทางชีววิทยา ซึ่งถ้าจะพูดถึงสถานภาพในอดีตเมื่อไม่กี่ปีมานี้ก็ได้มีสิ่งคล้าย ๆ กันที่พอจะเห็นได้ เช่น **โปรแกรมที่เป็น cellular automata** ที่สามารถสร้างแพตเทอร์นที่ซับซ้อนและสวยงาม โดยใช้ความรู้ทางคำนวณขั้นต้นในการควบคุมพฤติภาพของจุดของแสงบนจอภาพ (pixels) นับเป็นล้านจุดในการสร้างรูปหรือภาพต่าง ๆ จนกระทั่งในปัจจุบันที่คอมพิวเตอร์น่าจะสามารถติดตามพฤติภาพของเซลล์นับล้านได้และถ้าเป็นไปได้จริง นักวิทยาศาสตร์ก็จะสามารถติดตามกระบวนการที่เกี่ยวกับชีวิตทุกกระบวนการ ดังเช่นกรณีของ nematode ซึ่งเป็นหนึ่งในสิ่งมีชีวิตที่มีความเรียบง่ายที่สุด

เมื่อสามารถเริ่มต้นจาก nematode ได้ ก็น่าจะก้าวไปสู่เรื่องของการสร้างชีวิตในระดับทารกในครรภ์ได้ ซึ่งในขณะนี้นักวิทยาศาสตร์ (เช่นในสหรัฐฯ) สามารถมีทั้งความรู้และสรรพกำลังด้านคอมพิวเตอร์ที่จะติดตามปฏิกิริยาทั้งหมดในระดับเซลล์ อันถือเป็นองค์กรของข้อมูลที่เรียกว่า **อินโฟมอร์ฟ** (infomorph) ที่อยู่ในคอมพิวเตอร์ แม้ว่าในร่างกายของมนุษย์จะมีเซลล์นับเป็นจำนวนล้านล้าน (trillions of cells) ก็ตาม แต่สำหรับในแง่ของการทำความเข้าใจก็จะต้องมีการลดความซับซ้อนของปัญหาด้วย**การศึกษาเฉพาะ**สิ่งที่เห็นว่ามีความสำคัญที่สุดในร่างกายของมนุษย์ ซึ่งก็คงหนีไม่พ้นเรื่องของ "สมอง" อันหมายถึงว่าเรา**ไม่**จำเป็นต้องรู้ตำแหน่งและหน้าที่ของเซลล์ทุกเซลล์เพื่อให้ได้ชื่อว่าสามารถสร้างแบบจำลองที่เป็นตัวแทนของรูปร่างหน้าตา ความรู้สึก และพฤติภาพของมนุษย์คนหนึ่งในระดับ macro แต่สามารถที่จะทำการทดลองที่มีความซับซ้อนน้อยกว่านี้ได้หลายเท่า โดยยังคงรักษานัยสำคัญของความเป็นมนุษย์ไว้ได้โดยอาศัยแค่เพียงตัวอย่างของ DNA ของมนุษย์คนใดคนหนึ่งเพื่อใช้เป็นข้อมูลที่เป็นจุดเริ่มต้น โดยข้อมูลเช่นนี้อาจได้มาจากฐลีผงที่มาจากเส้นผม หรือเล็บนิ้วมือของใครคนใดคนหนึ่งที่เราสนใจหรือเคยมีความผูกพัน มาสกัดเอา DNA ออกมา เพื่อใช้สำหรับเป็นเมล็ดพันธุ์ในการสร้างมนุษย์เทียมขึ้นมาในคอมพิวเตอร์ โดยไม่ขัดต่อศีลธรรม ไม่ผิดกฎหมาย และไม่ก่อให้เกิดข้อพิพาทโต้แย้ง ซึ่งจะตามมาในภายหลัง

สมมติว่าท่านผู้อ่านอยากได้**อินโฟมอร์ฟของอัลเบิร์ต ไอน์สไตน์** (Albert Einstein) มาอยู่ในฮาร์ดไดรฟ์เพื่อทำการ**ซิมูเลต**ในเรื่องของปรีชาญาณของอัจฉริยะบุคคลของนักฟิสิกส์ผู้นี้ มันสมองของไอน์สไตน์ยังคงมีการเก็บรักษาไว้อยู่ในขณะนี้ แต่การที่จะเข้าถึงความมีปรีชาญาณของคนผู้นี้ก็ไม่จำเป็นต้องไปผ่าตัดมันสมองของท่านแต่อย่างใด สิ่งที่จำเป็นในการเรียกวิญญาณของนักปราชญ์ผู้นี้ให้คืนมาก็คือตัวอย่างของ DNA ที่ได้จากมันสมองของท่านผู้นี้เพื่อที่จะนำไปใช้ในการสร้างไอน์สไตน์เทียมขึ้นมาในคอมพิวเตอร์ แม้ว่าอินโฟมอร์ฟของท่านไอน์สไตน์ที่ได้มานั้นจะต้องได้รับการบำรุงเลี้ยงและให้การศึกษา เหมือน ๆ กับการเลี้ยงดูมนุษย์คนหนึ่งนับตั้งแต่เล็กจนเติบใหญ่ก็ตาม แต่ทันทีที่มนุษย์เทียมเช่นไอน์สไตน์ที่ถูกถอดแบบซึ่งได้เติบโตจนสู่วัยฉกรรจ์ด้วยอานุภาพของคอมพิวเตอร์เทคโนโลยี จนสามารถคิดได้เหมือนอัจฉริยบุรุษ ก็จะกลายเป็นสินค้าที่มีราคามากที่สุดอย่างหนึ่ง เพราะถ้าใครได้ไปก็จะได้ชื่อว่ามีมิตรที่มีความปราดเปรื่องแบบสุด ๆ ในทางคำนวณและในการมองจักรวาลที่หาตัวจับได้ยากในโลกไปอยู่คู่กาย โดยไปทำหน้าที่จุดประกายทางความคิดเช่นในแง่ของการสอนให้มองโลกหรือมองสิ่งที่กว้างไกลออกไป ในทัศนะของนักคำนวณซึ่งมีผลงานที่โลกยอมรับ ซึ่งต่อไปในอนาคตก็อาจจะมีการเรียกวิญญาณของผู้ที่เชี่ยวชาญในด้านต่าง ๆ ทั้งในทางโลกและในทาง**ธรรม** เพื่อมาชี้แนะให้มนุษย์ชาติคิดและก้าวไปในทางที่ถูกที่ควรเพื่อหลีกเลี่ยงความรุนแรง

สรุป

สถานะภาพการรุดหน้าในโลกของคอมพิวเตอร์ในปัจจุบันไม่ได้ยืนอยู่อย่างโดด ๆ ในโลกของฮาร์ดแวร์และซอฟต์แวร์ในรูปแบบที่เป็นธรรมดา ๆ เหมือนที่แล้ว ๆ มา แต่จะเข้าไปผสมผสานเพื่อใช้ประโยชน์จากผลของการค้นคว้าในด้านวิทยาศาสตร์ชีวภาพ โดยเฉพาะที่เกี่ยวกับพฤติภาพความเป็นไปในตัวมนุษย์ เพื่อนำผลดังกล่าวมาทำให้ผู้ใช้คอมพิวเตอร์ได้รับทานบารมีชนิดไม่เคยปรากฏมาก่อนในประวัติของเทคโนโลยี